Supplementary material

Population intervention effects in observational studies to emulate target trial results: reconciling effects of improved sanitation on child growth

Elizabeth T. Rogawski McQuade, Jade Benjamin-Chung, Daniel Westreich, and Benjamin F. Arnold

Details of the population intervention effects

The population intervention effects (PIE) estimate the expected difference in population LAZ comparing the exposure distribution after intervention implementation to the observed exposure distribution. The referent comparator in this contrast is commonly denoted as the natural course, or the expected value of the outcome under the observed distribution of exposure and covariates. The natural course can be expressed as the following by applying the law of total expectation:

$$\mathrm{E}[\mathrm{Y}] = \sum_{w} \sum_{a} \mathrm{E}[\mathrm{Y}|\mathrm{A}=a, \mathrm{W}=w] \Pr(\mathrm{A}=a|\mathrm{W}=w) \Pr(\mathrm{W}=w)$$

where Y is LAZ, A is improved sanitation, and W are confounders.

This expression can then be modified to denote the expected value of the outcome under the intervention exposure distribution (i.e. the counterfactual scenario in the PIE contrast) by replacing the observed distribution of exposure, $\Pr(\mathrm{A}=a|\mathrm{W}=w)$, with the intervened-on distribution of exposure, $\Pr(\mathrm{g}(\mathrm{A})=a|\mathrm{W}=w)$, such that:

$$\mathrm{E}\left[\mathrm{Y}^{intervention}\right] = \sum_{w} \sum_{a} \mathrm{E}[\mathrm{Y}|\mathrm{A}=a, \mathrm{W}=w] \Pr(\mathrm{g}(\mathrm{A})=a|\mathrm{W}=w) \Pr(\mathrm{W}=w)$$

The PIE is therefore:

$$\begin{aligned}\varphi_{\mathrm{PIE}} &= \mathrm{E}\left[\mathrm{Y}^{intervention}\right] - \mathrm{E}[\mathrm{Y}] \\ &= \sum_{w} \sum_{a} \mathrm{E}[\mathrm{Y}|\mathrm{A}=a, \mathrm{W}=w] \Pr(\mathrm{g}(\mathrm{A})=a|\mathrm{W}=w) \Pr(\mathrm{W}=w) \\ &\quad - \mathrm{E}[\mathrm{Y}|\mathrm{A}=a, \mathrm{W}=w] \Pr(\mathrm{A}=a|\mathrm{W}=w) \Pr(\mathrm{W}=w)\end{aligned}$$

$$= \sum_{w} \sum_{a} [\mathrm{E}(\mathrm{Y}|\mathrm{A}=a, \mathrm{W}=w) * [\Pr(\mathrm{g}(\mathrm{A})=a\ |\mathrm{W}=w) - \Pr(\mathrm{A}=a|\mathrm{W}=w)] * \Pr(\mathrm{W}=w)]$$

For the intervention that deterministically provided improved sanitation to all children, the intervened-on distribution of exposure was:

$$\Pr(\mathrm{g}(\mathrm{A})=a|\mathrm{W}=w) = \Pr(\mathrm{g}(\mathrm{A})=1|\mathrm{W}=w) = 1$$

For the intervention that stochastically increased the prevalence of improved sanitation by x (e.g. 30 or 60 percentage points) from the observed prevalence, individuals observed to have improved sanitation (A=1) continued to have improved sanitation:

$$\Pr(\mathrm{g}(\mathrm{A})=1|\mathrm{A}=1, \mathrm{W}=w) = 1$$

and individuals observed to not have improved sanitation (A=0) were randomly assigned improved sanitation at a probability such that the overall prevalence increased by x:

$$\Pr(\mathrm{g}(\mathrm{A})=1|\mathrm{A}=0, \mathrm{W}=w) = \frac{\min(0.9, \Pr(\mathrm{A}=1|\mathrm{W}=w)+x) - \Pr(\mathrm{A}=1|\mathrm{W}=w)}{1-\Pr(\mathrm{A}=1|\mathrm{W}=w)}$$

The increase was limited to not exceed 90%, which was only relevant for populations in which the observed prevalence of improved sanitation exceeded ($0.9 - x$). These PIE were not estimated in populations with an observed prevalence of 90% or greater.

The overall intervened-on distribution of exposure was therefore the weighted average of these two expressions:

$$\Pr(\mathrm{g(A)} = 1|\mathrm{W} = w)$$

$$= \Pr(\mathrm{A} = 1|\mathrm{W} = w) + \frac{\min(0.9, \Pr(\mathrm{A}=1|\mathrm{W}=w)+x) - \Pr(\mathrm{A}=1|\mathrm{W}=w)}{1-\Pr(\mathrm{A}=1|\mathrm{W}=w)} * [1 - \Pr(\mathrm{A} = 1|\mathrm{W} = w)]$$

$$= \max(0.9, \Pr(\mathrm{A} = 1|\mathrm{W} = w) + x)$$

In all cases, for the binary exposure:

$$\Pr(\mathrm{g(A)} = 0|\mathrm{W} = w) = 1 - \Pr(\mathrm{g(A)} = 1|\mathrm{W} = w)$$

For the interventions that stochastically increased the prevalence of improved sanitation by x, the assumption of conditional exchangeability did not apply since the intervention, and therefore the associated expected value of LAZ under the intervention, depended on the observed exposure. In these cases, the appropriate alternative assumption is described by Young et al. in Appendix B,[1] briefly, that there are no open backdoor paths between the exposure and the potential outcome in the Single World Intervention Graph (SWIG) for the intervention.

Derivation of the relationship between the average treatment effect (ATE) and population intervention effect (PIE)

The PIE is directly proportional to the ATE by x, the increase the prevalence of improved sanitation from the baseline observed exposure achieved by the intervention. This relationship can be derived from the expressions for the ATE and PIE by the following:

Starting with the expression for the PIE:

$$\varphi_{\mathrm{PIE}} = \sum_{w} \sum_{a} [\mathrm{E}(\mathrm{Y}|\mathrm{A}=a, \mathrm{W}=w) * [\Pr(\mathrm{g}(\mathrm{A})=a \,|\mathrm{W}=w) \; - \; \Pr(\mathrm{A}=a|\mathrm{W}=w)] * \Pr(\mathrm{W}=w)]$$

Expanding the summation over a:

$$\varphi_{\mathrm{PIE}} = \sum_{w} [\{\mathrm{E}(\mathrm{Y}|\mathrm{A}=1, \mathrm{W}=w) * [\Pr(\mathrm{g}(\mathrm{A})=1 \,|\mathrm{W}=w) - \Pr(\mathrm{A}=1|\mathrm{W}=w)]$$

$$+ [\mathrm{E}(\mathrm{Y}|\mathrm{A}=0, \mathrm{W}=w) * [\Pr(\mathrm{g}(\mathrm{A})=0 \,|\mathrm{W}=w) - \Pr(\mathrm{A}=0|\mathrm{W}=w)]\} * \Pr(\mathrm{W}=w)\}]$$

Replacing the probabilities for $\mathrm{g}(\mathrm{A})=0$ and $\mathrm{A}=0$ with the inverse of those for $\mathrm{g}(\mathrm{A})=1$ and $\mathrm{A}=1$:

$$\varphi_{\mathrm{PIE}} = \sum_{w} [\{\mathrm{E}(\mathrm{Y}|\mathrm{A}=1, \mathrm{W}=w) * [\Pr(\mathrm{g}(\mathrm{A})=1 \,|\mathrm{W}=w) - \Pr(\mathrm{A}=1|\mathrm{W}=w)]$$

$$+ [\mathrm{E}(\mathrm{Y}|\mathrm{A}=0, \mathrm{W}=w) * [(1 - \Pr(\mathrm{g}(\mathrm{A})=1 \,|\mathrm{W}=w)) - (1 - \Pr(\mathrm{A}=1|\mathrm{W}=w))]\} * \Pr(\mathrm{W}=w)]$$

Simplifying and rearranging the equation:

$$\varphi_{\mathrm{PIE}} = \sum_{w} [\{\mathrm{E}(\mathrm{Y}|\mathrm{A}=1, \mathrm{W}=w) * [\Pr(\mathrm{g}(\mathrm{A})=1 \,|\mathrm{W}=w) - \Pr(\mathrm{A}=1|\mathrm{W}=w)]$$

$$- [\mathrm{E}(\mathrm{Y}|\mathrm{A}=0, \mathrm{W}=w) * [\Pr(\mathrm{g}(\mathrm{A})=1 \,|\mathrm{W}=w) - \Pr(\mathrm{A}=1|\mathrm{W}=w)]\} * \Pr(\mathrm{W}=w)]$$

$$\varphi_{\mathrm{PIE}} = \sum_{w} [\{\mathrm{E}(\mathrm{Y}|\mathrm{A}=1, \mathrm{W}=w) - \; [\mathrm{E}(\mathrm{Y}|\mathrm{A}=0, \mathrm{W}=w)]\}$$

$$* [\Pr(\mathrm{g}(\mathrm{A})=1 \,|\mathrm{W}=w) - \Pr(\mathrm{A}=1|\mathrm{W}=w)] * \Pr(\mathrm{W}=w)]$$

Factoring out the ATE:

$$\varphi_{\mathrm{PIE}} = \varphi_{\mathrm{ATE}} * [\Pr(\mathrm{g}(\mathrm{A})=1 \,|\mathrm{W}=w) - \Pr(\mathrm{A}=1|\mathrm{W}=w)]$$

Plugging in the expression for $\Pr(\mathrm{g}(\mathrm{A})=1 \,|\mathrm{W}=w)$:

$$\varphi_{\mathrm{PIE}} = \varphi_{\mathrm{ATE}} * [\Pr(\mathrm{A}=1|\mathrm{W}=w) + x - \Pr(\mathrm{A}=1|\mathrm{W}=w)]$$

$$= \varphi_{\mathrm{ATE}} * x$$

Figure S1. Schematic describing the contrasts for the three target parameters: 100% compared to 0% improved sanitation coverage in non-exchangeable populations for the unadjusted average treatment effect (uATE), 100% compared to 0% coverage in exchangeable populations for the average treatment effect (ATE), and increased coverage (e.g. 100%, 80%, or 50%) compared to baseline observed coverage (e.g. 20%) in exchangeable populations for the population intervention effects (PIEs). The ATE and PIE assume no unmeasured confounding. Diagrams adapted from.[2]

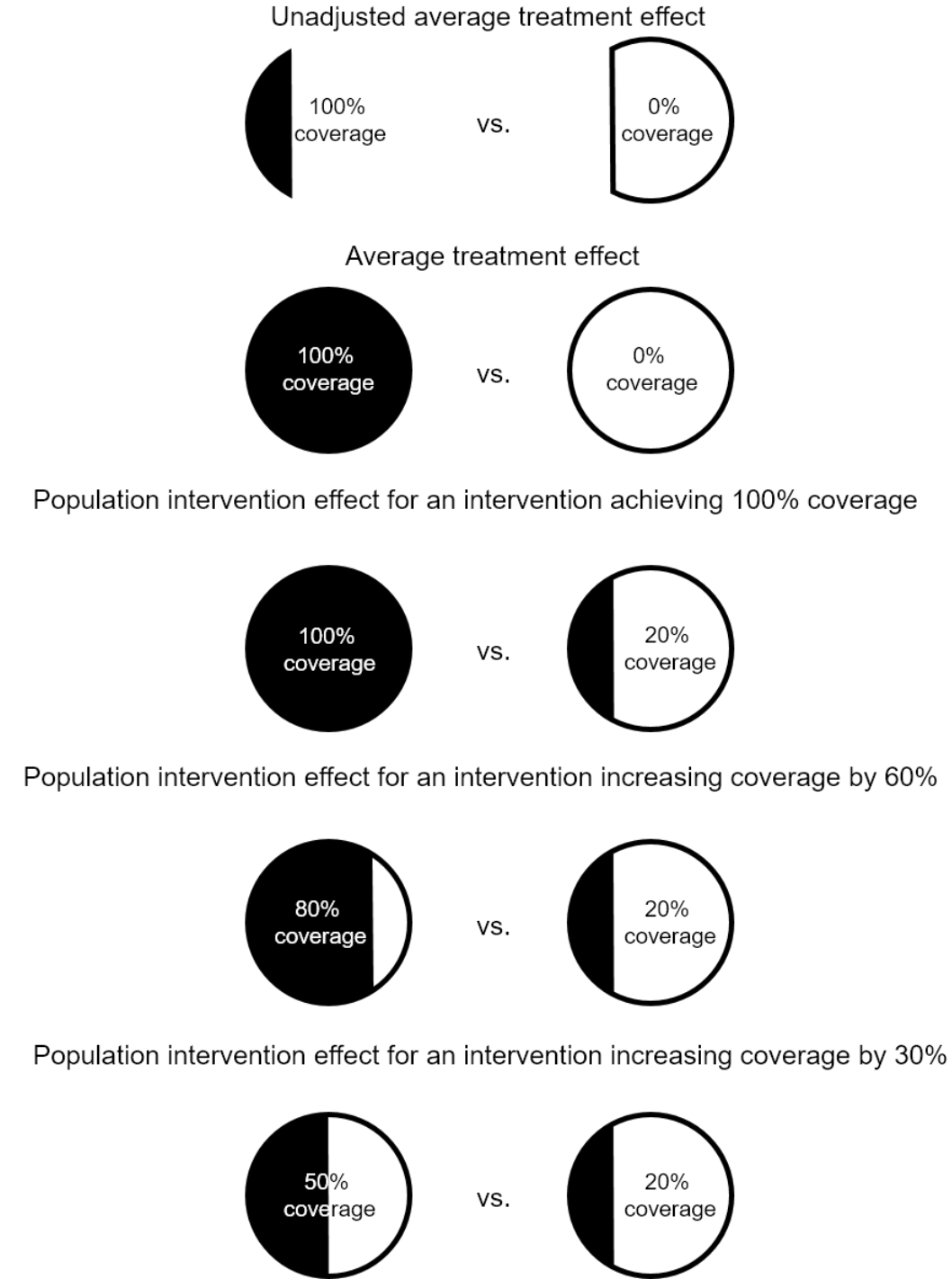


Table S1. Unadjusted average treatment effect, adjusted average treatment effect, and population intervention effect for an intervention achieving 100% improved sanitation coverage in each Demographic and Health Survey (DHS).

Region	Country	Sanitation coverage* (%)	Unadjusted average treatment effect (95% CI)	Adjusted average treatment effect (95% CI)	Population intervention (100% coverage) effect (95% CI)
Western Africa	Senegal	59.7	0.29 (0.19, 0.40)	-0.03 (-0.14, 0.07)	-0.02 (-0.07, 0.03)
Western Africa	Nigeria	42.2	0.30 (0.17, 0.43)	0.06 (-0.06, 0.18)	0.03 (-0.03, 0.11)
Western Africa	Gambia	37.0	0.16 (-0.08, 0.40)	0.07 (-0.14, 0.28)	0.05 (-0.09, 0.19)
Western Africa	Burkina Faso	36.4	0.25 (0.00, 0.50)	0.21 (-0.03, 0.45)	0.13 (-0.02, 0.28)
Western Africa	Niger	34.7	-0.02 (-0.36, 0.31)	-0.20 (-0.53, 0.15)	-0.12 (-0.32, 0.09)
Western Africa	Mali	31.8	0.11 (-0.03, 0.27)	-0.09 (-0.22, 0.07)	-0.06 (-0.15, 0.05)
Western Africa	Togo	26.3	0.03 (-0.23, 0.34)	-0.03 (-0.27, 0.24)	-0.02 (-0.21, 0.19)
Western Africa	Guinea	25.7	0.26 (-0.03, 0.53)	0.27 (-0.03, 0.53)	0.20 (-0.02, 0.39)
Western Africa	Cote d'Ivoire	24.8	0.35 (0.07, 0.63)	0.40 (0.10, 0.70)	0.30 (0.07, 0.53)
Western Africa	Benin	24.3	0.28 (0.16, 0.42)	0.07 (-0.06, 0.21)	0.05 (-0.05, 0.16)
Western Africa	Liberia	23.8	0.21 (-0.30, 0.70)	0.11 (-0.39, 0.59)	0.09 (-0.31, 0.48)
Western Africa	Ghana	14.6	0.34 (-0.07, 0.70)	0.12 (-0.18, 0.40)	0.10 (-0.16, 0.34)
Western Africa	Sierra Leone	9.6	0.19 (-0.33, 0.76)	-0.02 (-0.55, 0.64)	-0.02 (-0.50, 0.58)
Eastern Africa	Rwanda	54.8	0.09 (-0.08, 0.25)	0.02 (-0.15, 0.16)	0.01 (-0.07, 0.07)
Eastern Africa	Malawi	52.5	-0.01 (-0.16, 0.15)	-0.02 (-0.17, 0.15)	-0.01 (-0.08, 0.07)
Eastern Africa	Burundi	43.6	0.19 (0.08, 0.31)	0.07 (-0.03, 0.18)	0.04 (-0.02, 0.09)
Eastern Africa	Zimbabwe	41.6	0.20 (0.00, 0.38)	0.14 (-0.06, 0.33)	0.08 (-0.03, 0.18)
Eastern Africa	Zambia	32.2	0.11 (-0.03, 0.23)	-0.02 (-0.17, 0.12)	-0.01 (-0.11, 0.08)
Eastern Africa	Mozambique	30.8	0.25 (0.06, 0.44)	0.02 (-0.16, 0.22)	0.02 (-0.11, 0.15)
Eastern Africa	Comoros	25.9	0.14 (-0.32, 0.58)	-0.01 (-0.52, 0.43)	-0.01 (-0.37, 0.32)
Eastern Africa	Kenya	23.9	0.24 (0.11, 0.37)	0.03 (-0.10, 0.15)	0.03 (-0.07, 0.11)
Eastern Africa	Uganda	19.7	0.21 (-0.01, 0.44)	0.02 (-0.21, 0.25)	0.01 (-0.17, 0.21)
Eastern Africa	Tanzania	17.6	0.37 (0.20, 0.54)	0.10 (-0.07, 0.28)	0.08 (-0.05, 0.22)
Eastern Africa	Ethiopia	7.8	0.57 (0.22, 0.86)	0.29 (-0.03, 0.58)	0.26 (-0.02, 0.50)
Eastern Africa	Madagascar	3.9	0.02 (-0.56, 0.82)	-0.48 (-1.11, 0.30)	-0.45 (-1.04, 0.28)
Middle/Southern Africa	Sao Tome and Principe	74.4	-0.66 (-1.58, 0.14)	-0.83 (-1.63, -0.05)	-0.21 (-0.42, -0.01)
Middle/Southern Africa	Lesotho	69.6	-0.08 (-0.40, 0.28)	0.07 (-0.33, 0.51)	0.02 (-0.10, 0.14)
Middle/Southern Africa	Eswatini	65.4	-0.05 (-0.34, 0.24)	-0.02 (-0.31, 0.28)	-0.01 (-0.10, 0.09)
Middle/Southern Africa	South Africa	61.0	-0.12 (-0.53, 0.30)	-0.33 (-0.83, 0.13)	-0.14 (-0.34, 0.05)
Middle/Southern Africa	Namibia	57.0	0.04 (-0.36, 0.45)	-0.00 (-0.41, 0.40)	-0.00 (-0.16, 0.16)
Middle/Southern Africa	Angola	51.1	0.43 (0.22, 0.65)	0.20 (-0.00, 0.43)	0.11 (-0.00, 0.23)
Middle/Southern Africa	Cameroon	38.0	0.13 (-0.08, 0.35)	-0.06 (-0.27, 0.17)	-0.04 (-0.17, 0.11)
Middle/Southern Africa	Gabon	33.7	0.52 (0.28, 0.75)	0.05 (-0.23, 0.37)	0.04 (-0.19, 0.29)
Middle/Southern Africa	Chad	22.0	0.30 (-0.04, 0.63)	0.29 (0.01, 0.58)	0.22 (0.01, 0.44)
Middle/Southern Africa	Congo Democratic Republic	19.9	0.06 (-0.23, 0.31)	0.09 (-0.17, 0.34)	0.07 (-0.14, 0.27)

Middle/Southern Africa	Congo	9.5	0.15 (-0.19, 0.50)	0.04 (-0.29, 0.37)	0.04 (-0.27, 0.34)
South & Southeast Asia	Maldives	98.3	0.53 (-0.31, 1.41)	0.66 (-0.11, 1.59)	0.01 (-0.00, 0.03)
South & Southeast Asia	Pakistan	77.7	0.56 (0.32, 0.80)	0.31 (0.05, 0.56)	0.06 (0.01, 0.11)
South & Southeast Asia	India	76.4	0.20 (0.13, 0.27)	0.00 (-0.06, 0.07)	0.00 (-0.02, 0.02)
South & Southeast Asia	Cambodia	76.3	0.08 (-0.17, 0.32)	-0.12 (-0.37, 0.14)	-0.02 (-0.07, 0.03)
South & Southeast Asia	Nepal	74.3	-0.10 (-0.34, 0.14)	-0.09 (-0.31, 0.15)	-0.02 (-0.08, 0.04)
South & Southeast Asia	Timor-Leste	70.4	0.25 (-0.09, 0.58)	0.08 (-0.25, 0.44)	0.02 (-0.07, 0.13)
South & Southeast Asia	Myanmar	48.2	-0.01 (-0.20, 0.16)	-0.19 (-0.37, -0.01)	-0.09 (-0.17, -0.00)
South & Southeast Asia	Bangladesh	46.9	0.24 (0.14, 0.36)	0.01 (-0.10, 0.13)	0.01 (-0.05, 0.07)
Latin America & Caribbean	Peru	88.0	0.02 (-0.12, 0.16)	-0.02 (-0.14, 0.12)	-0.00 (-0.02, 0.02)
Latin America & Caribbean	Guatemala	80.9	0.03 (-0.07, 0.14)	0.00 (-0.10, 0.11)	0.00 (-0.02, 0.02)
Latin America & Caribbean	Dominican Republic	79.6	0.17 (-0.11, 0.43)	-0.04 (-0.28, 0.21)	-0.01 (-0.06, 0.05)
Latin America & Caribbean	Guyana	76.1	0.34 (-0.02, 0.76)	0.28 (-0.11, 0.70)	0.07 (-0.03, 0.19)
Latin America & Caribbean	Honduras	68.1	0.27 (0.15, 0.38)	0.01 (-0.09, 0.12)	0.00 (-0.03, 0.04)
Latin America & Caribbean	Haiti	38.8	0.11 (-0.07, 0.29)	-0.04 (-0.23, 0.16)	-0.02 (-0.15, 0.10)
Other	Tajikistan	97.6	-0.09 (-0.47, 0.30)	0.02 (-0.30, 0.41)	0.00 (-0.01, 0.01)
Other	Albania	96.6	0.95 (0.07, 1.83)	0.47 (-0.38, 1.30)	0.01 (-0.01, 0.04)
Other	Turkey	95.6	0.41 (-0.06, 0.87)	-0.12 (-0.59, 0.33)	-0.01 (-0.04, 0.02)
Other	Kyrgyz Republic	94.8	0.27 (-0.21, 0.78)	0.11 (-0.33, 0.63)	0.00 (-0.01, 0.03)
Other	Egypt	89.4	-0.40 (-0.73, -0.13)	-0.45 (-0.78, -0.17)	-0.04 (-0.06, -0.01)
Other	Armenia	79.7	0.21 (-0.28, 0.68)	0.12 (-0.54, 0.68)	0.02 (-0.11, 0.14)
Other	Azerbaijan	73.4	0.05 (-0.24, 0.34)	0.03 (-0.25, 0.30)	0.01 (-0.07, 0.08)
Other	Yemen	59.7	0.39 (0.24, 0.53)	0.13 (-0.01, 0.27)	0.05 (-0.00, 0.11)
Other	Papua New Guinea	25.4	0.51 (-0.01, 1.03)	-0.20 (-0.59, 0.19)	-0.14 (-0.42, 0.14)

*

Figure S2. Average treatment effects (ATE) and population intervention effects (PIE) for an intervention achieving 100% improved sanitation coverage with 95% confidence intervals for each study site by coverage of improved sanitation and mean length-for-age z-score (LAZ). Estimates are shaded by region and shaped by study design.

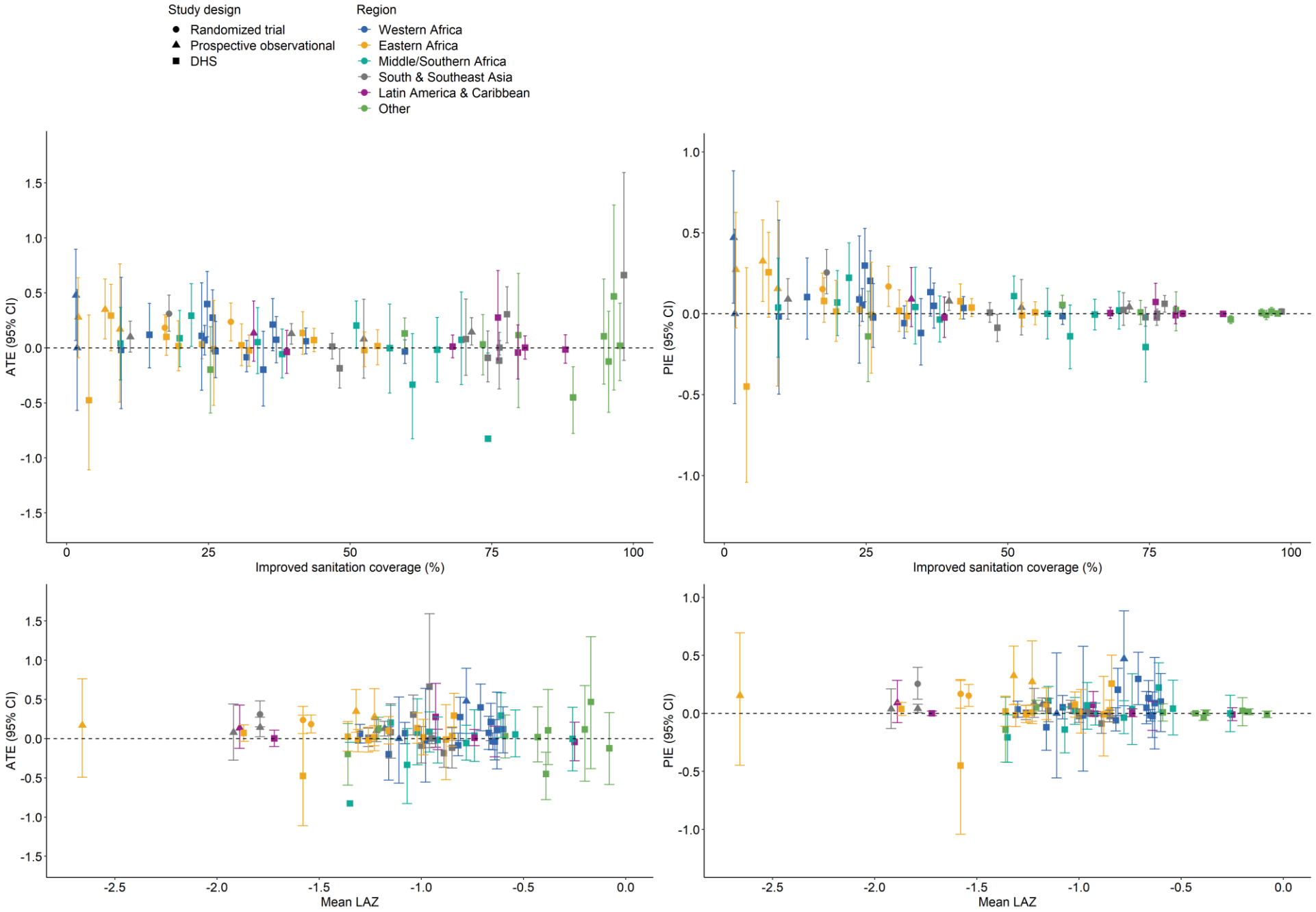
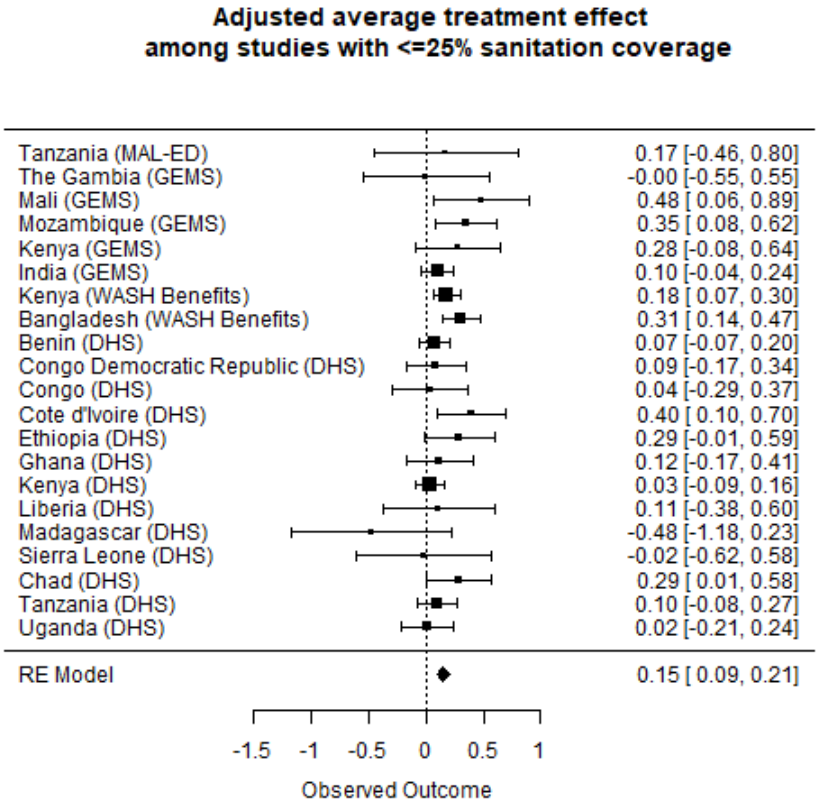
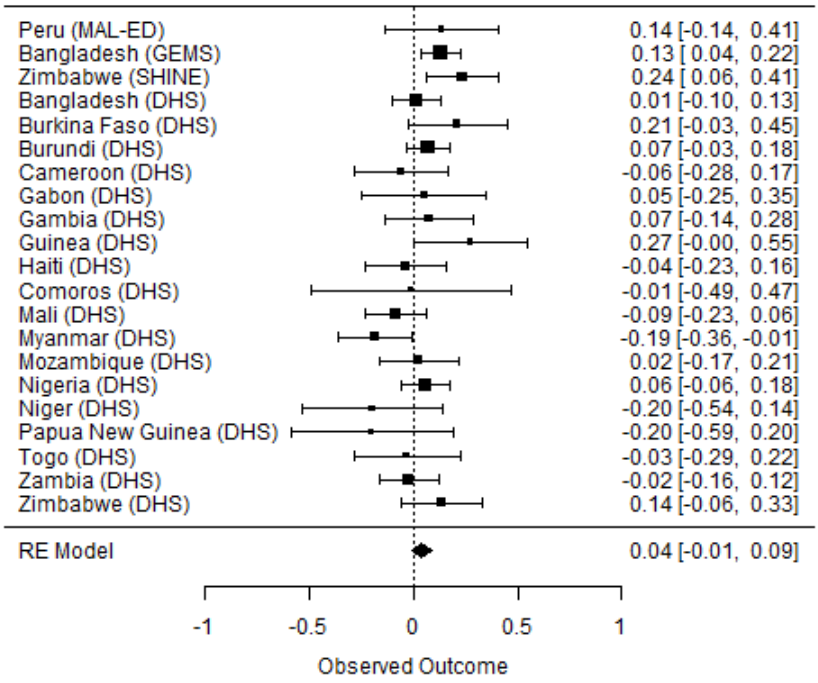


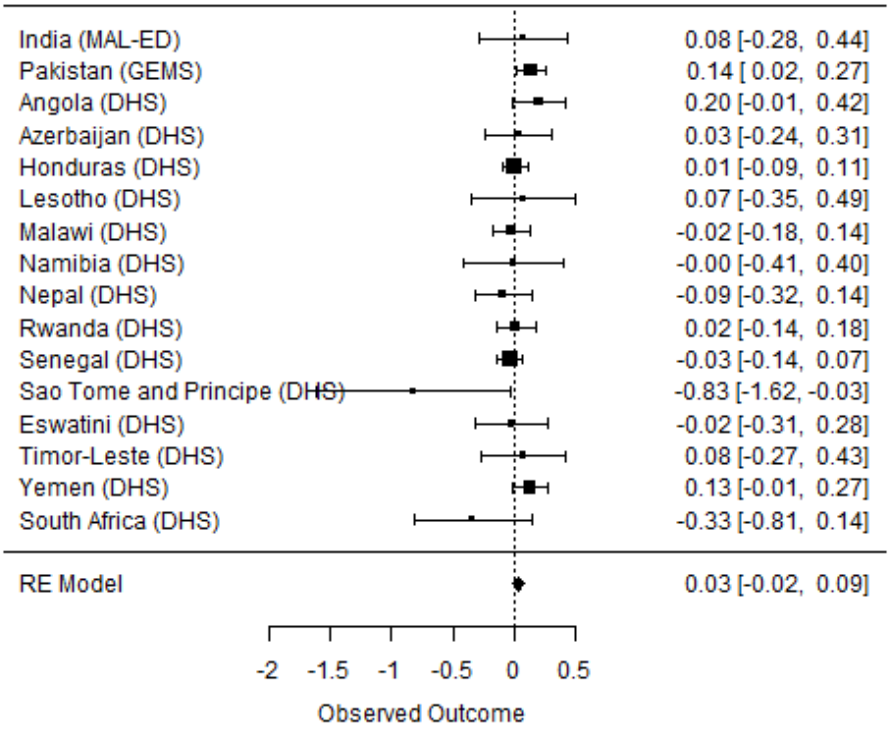
Figure S3. Forest plots for average treatment effects and population intervention effects for an intervention achieving 100% improved sanitation coverage by coverage of improved sanitation and mean length-for-age z-score (LAZ).



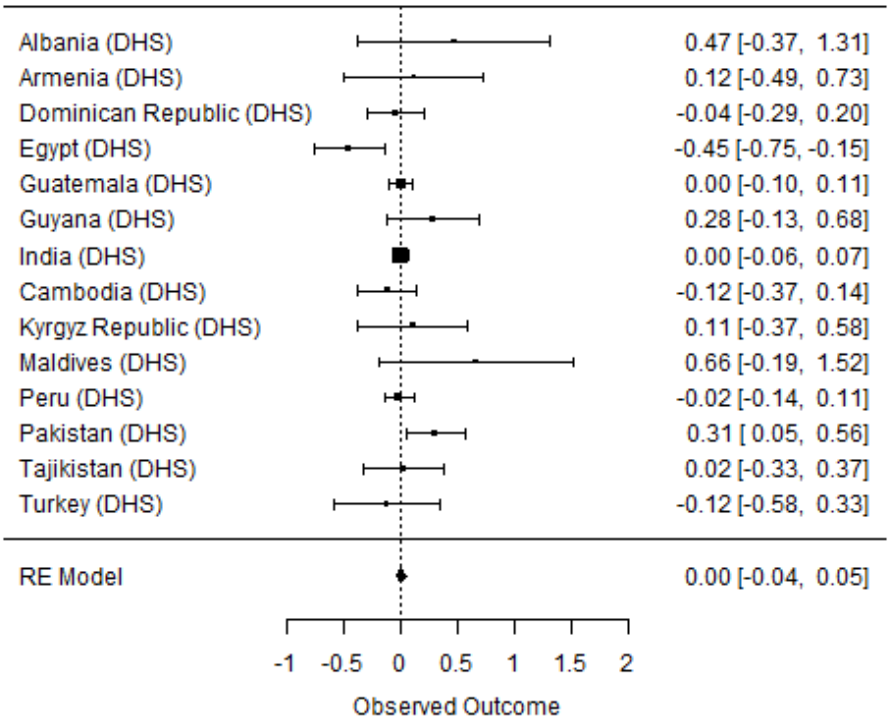
Adjusted average treatment effect among studies with 26-50% sanitation coverage



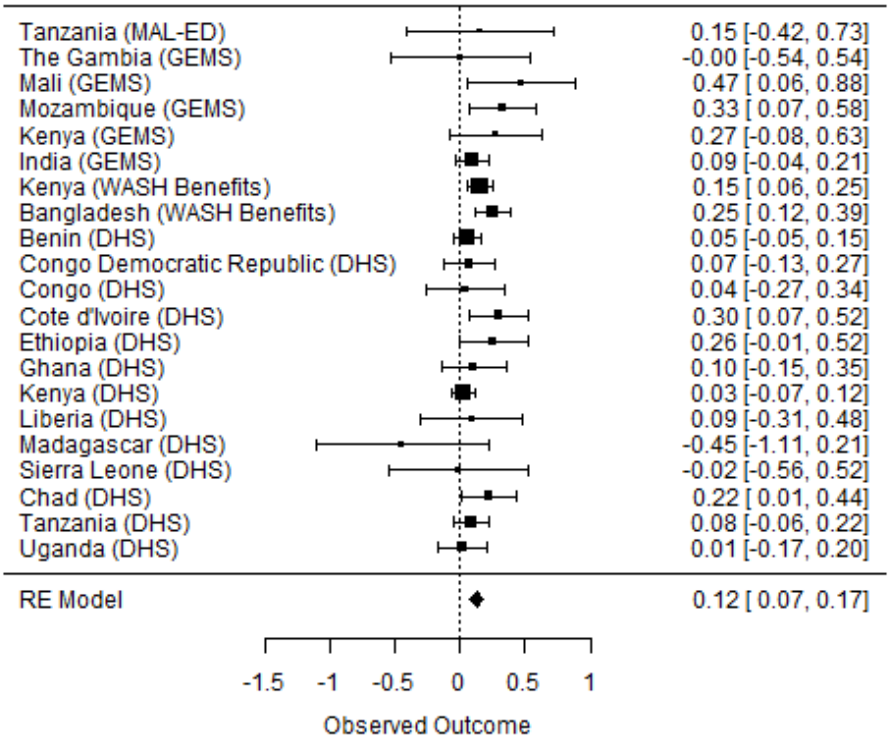
Adjusted average treatment effect among studies with 51-75% sanitation coverage



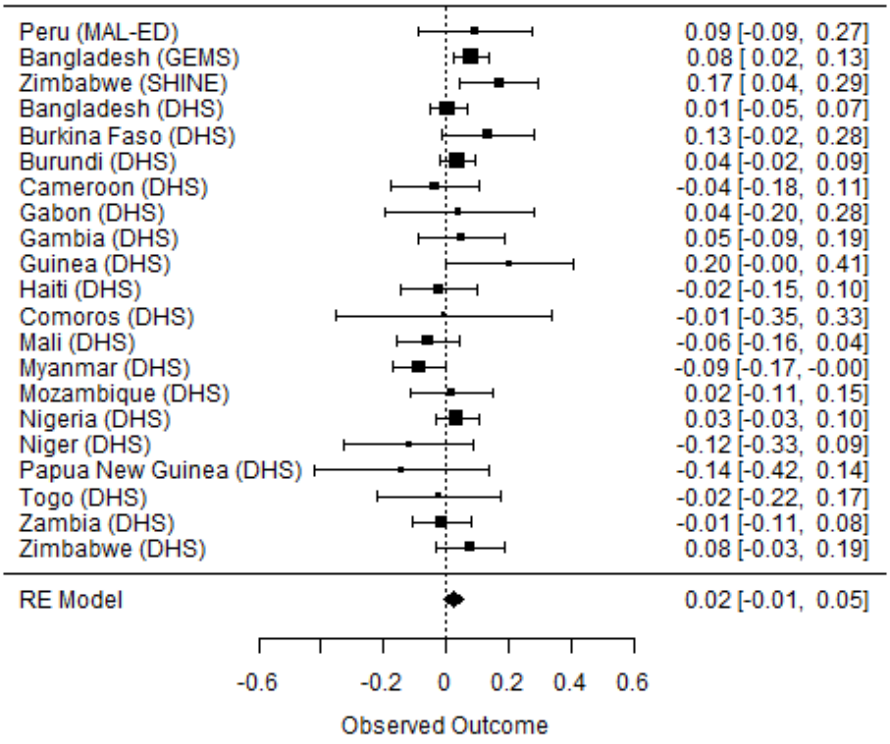
Adjusted average treatment effect among studies with 76+% sanitation coverage



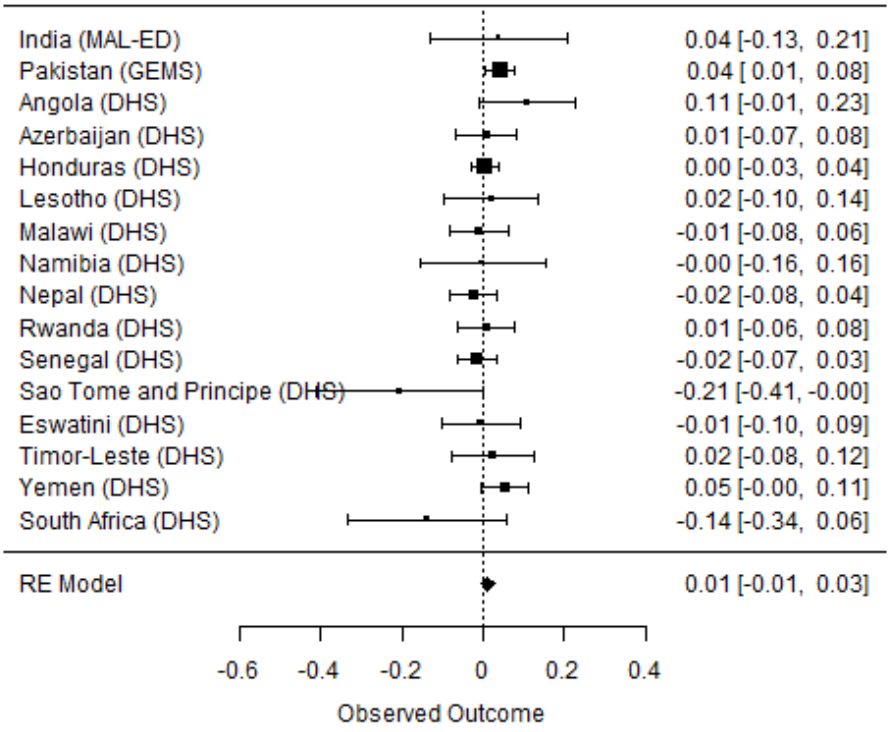
Population intervention effect
among studies with <=25% sanitation coverage



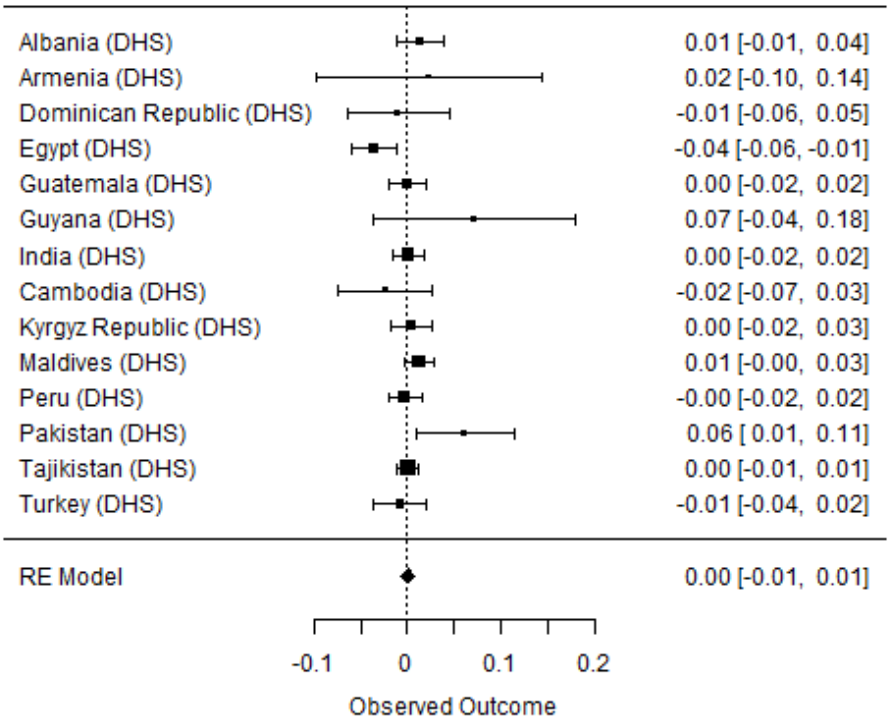
Population intervention effect
among studies with 26-50% sanitation coverage



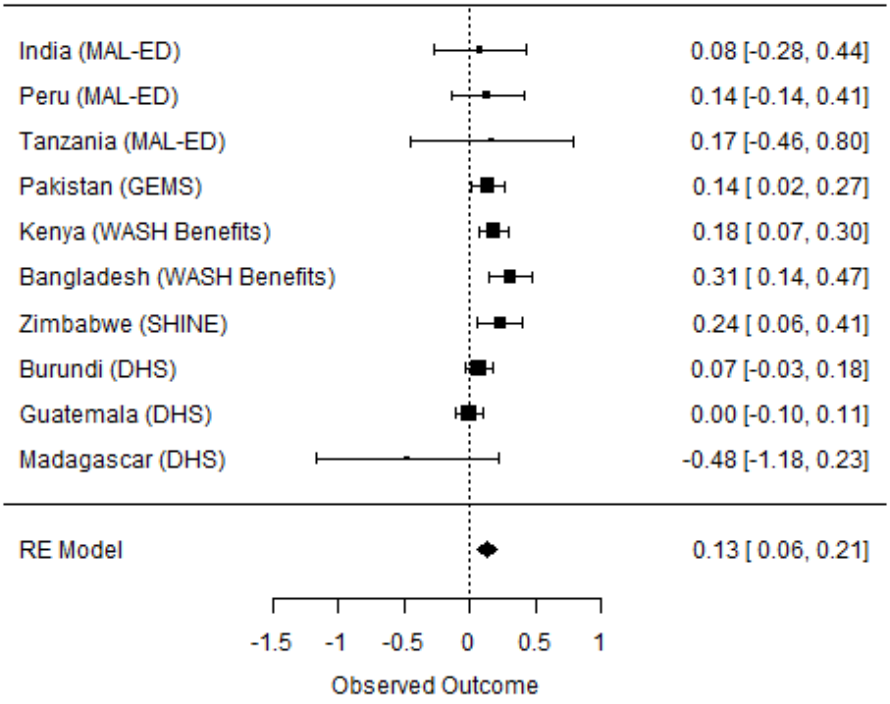
Population intervention effect among studies with 51-75% sanitation coverage



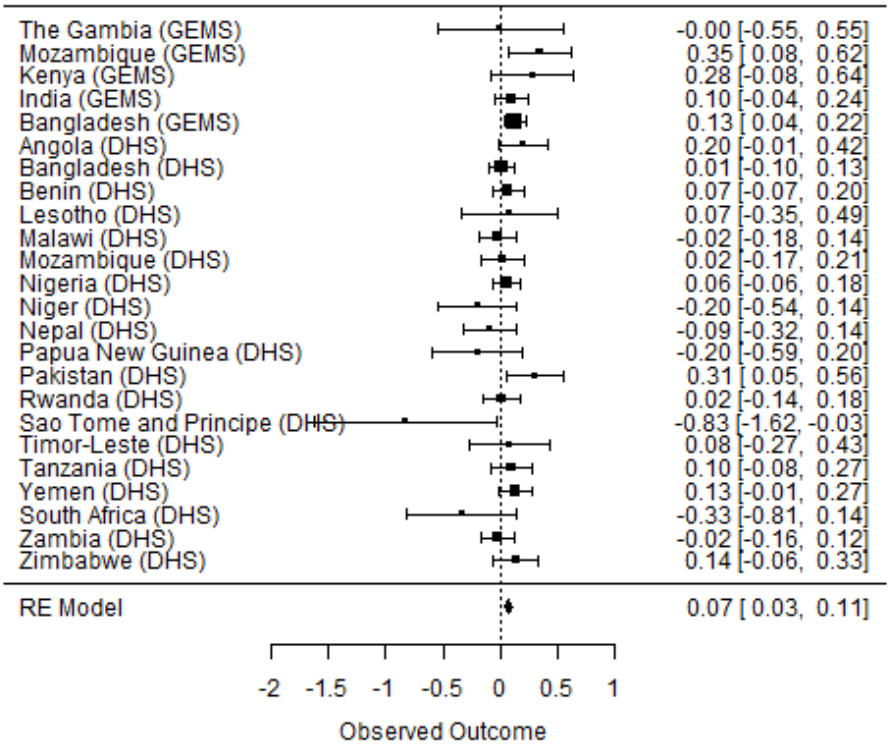
Population intervention effect among studies with 76+% sanitation coverage



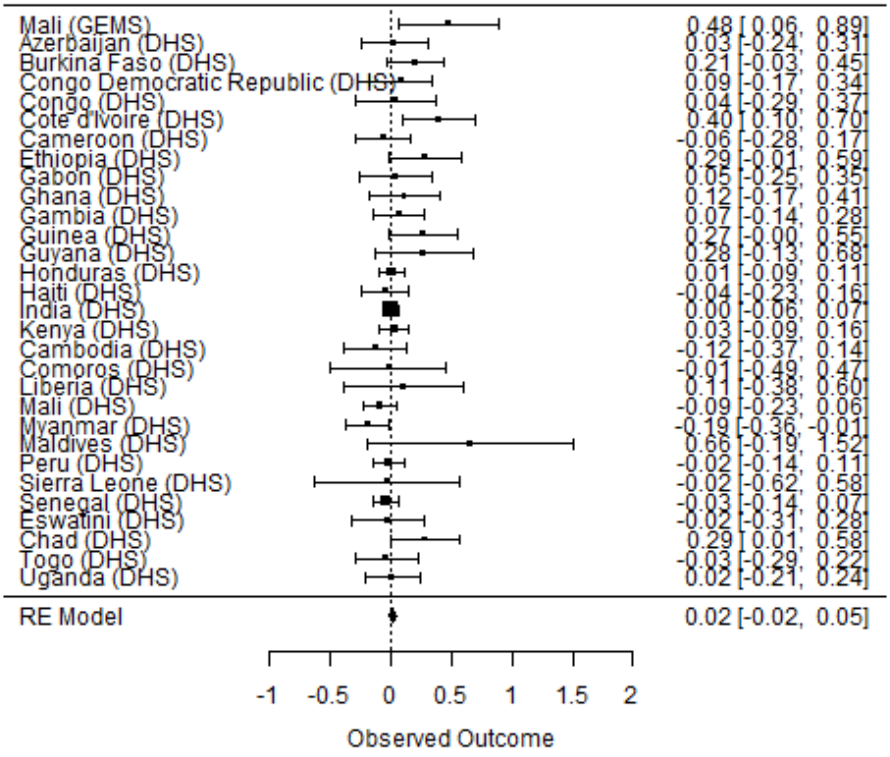
Adjusted average treatment effect among studies with <-1.5 mean LAZ



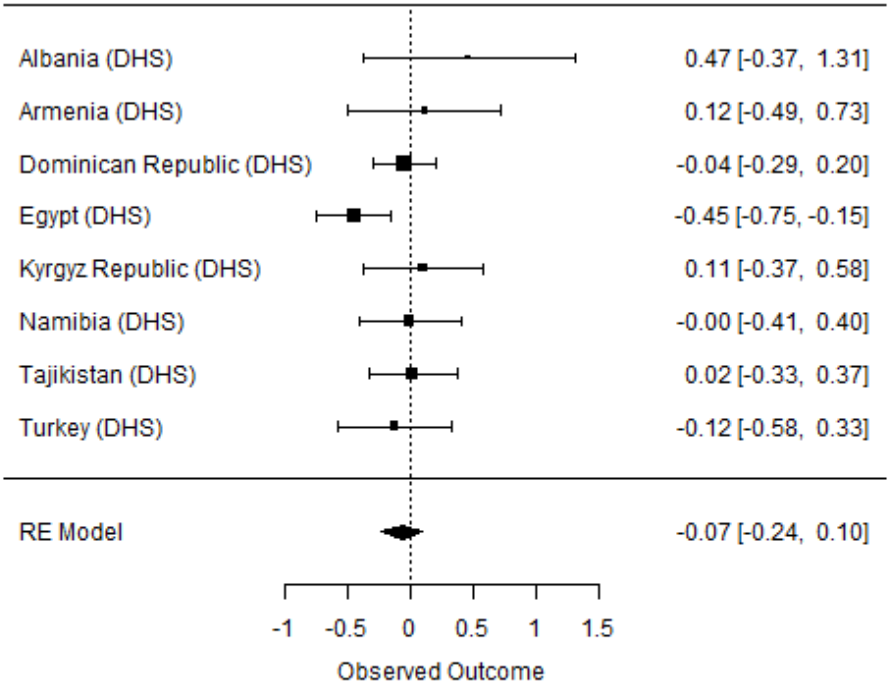
Adjusted average treatment effect among studies with -1.5 to -1 mean LAZ



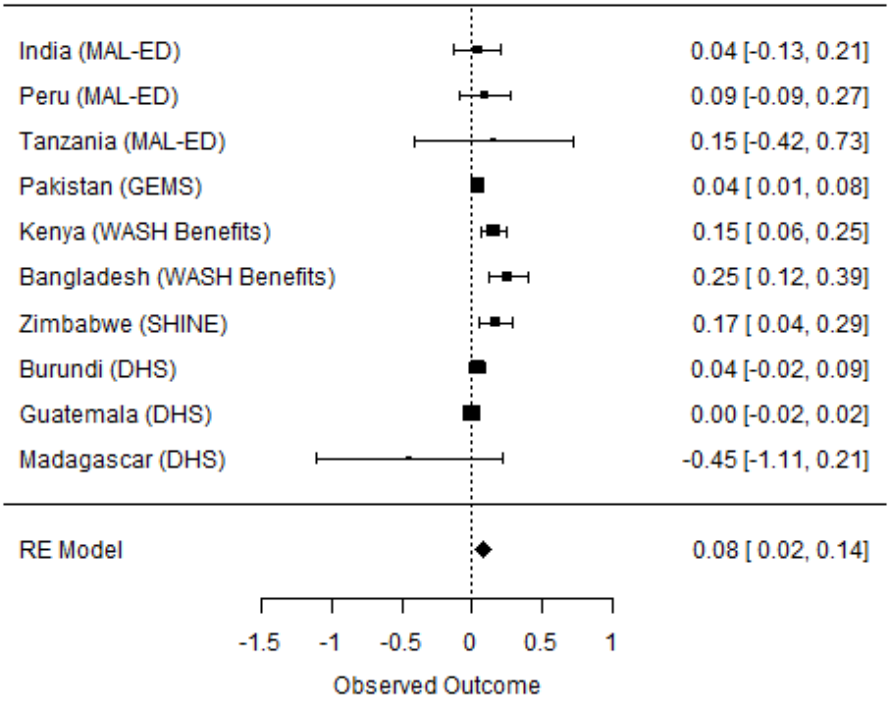
Adjusted average treatment effect among studies with -1 to -0.5 mean LAZ



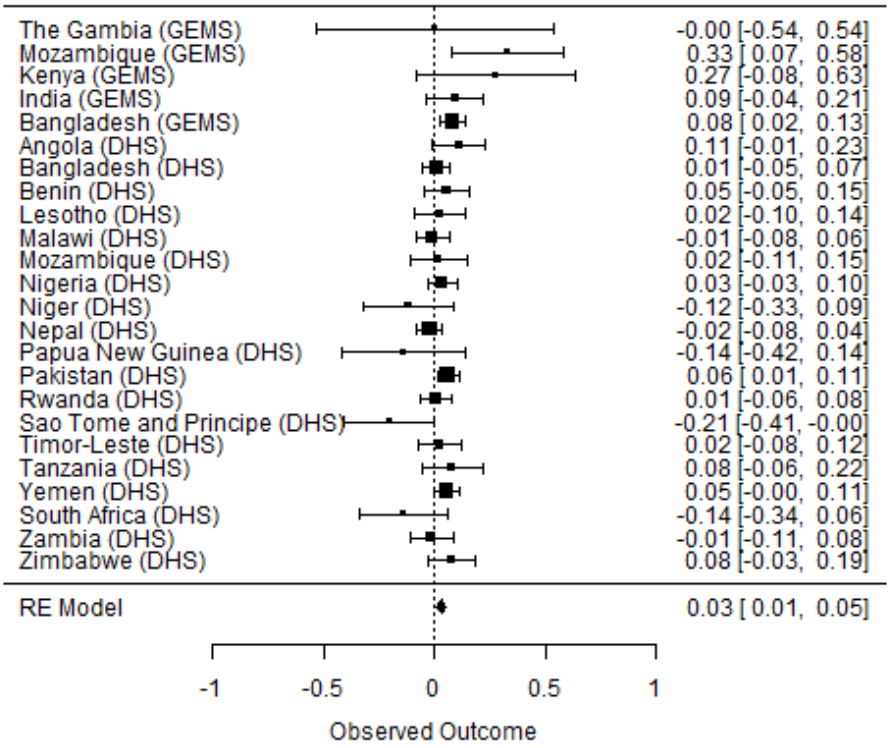
Adjusted average treatment effect among studies with -0.5+ mean LAZ



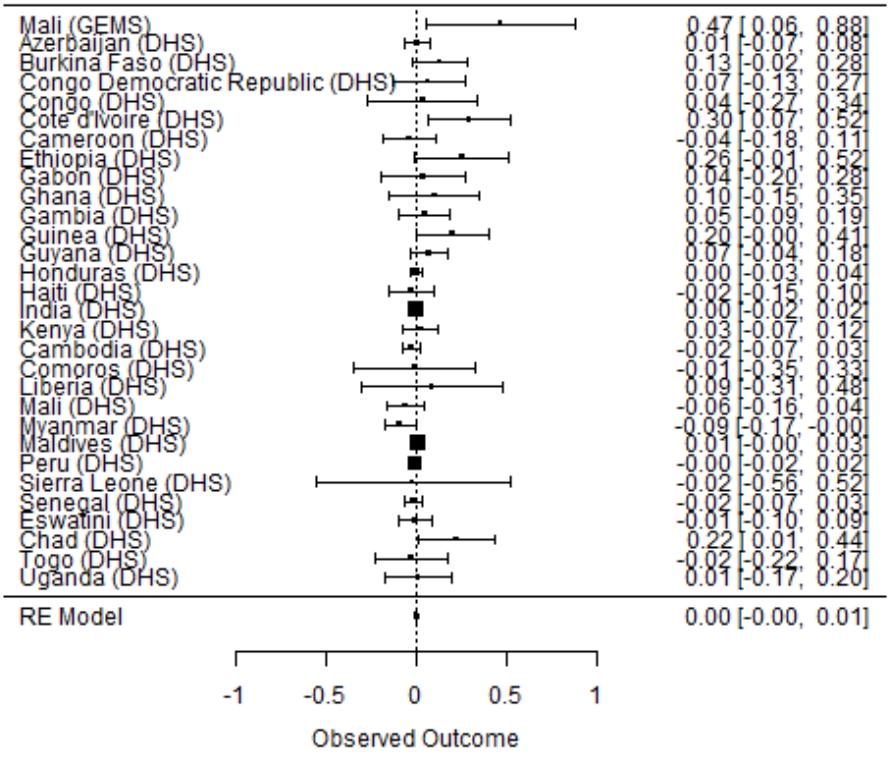
Population intervention effect among studies with <-1.5 mean LAZ



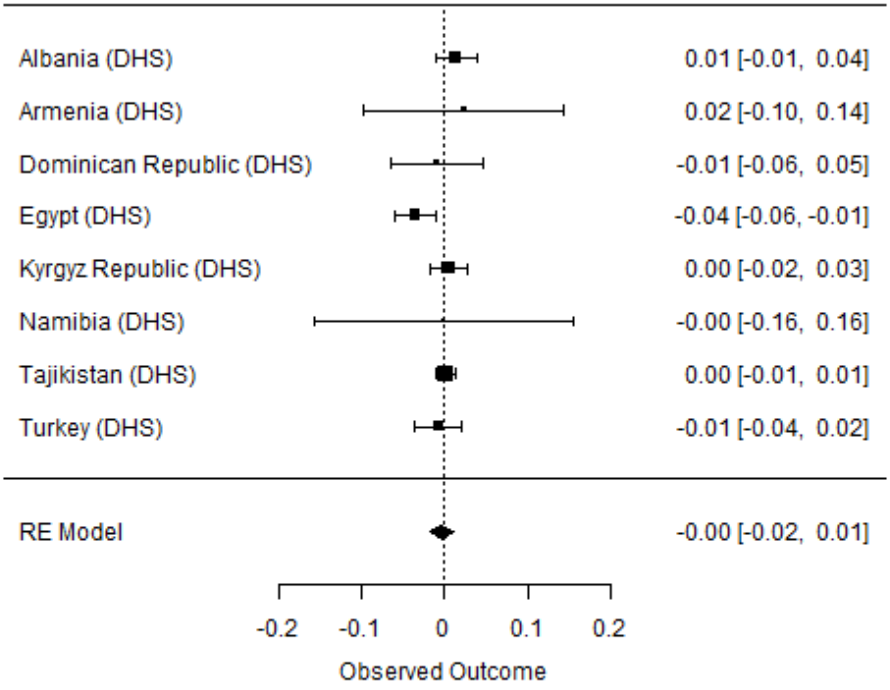
Population intervention effect among studies with -1.5 to -1 mean LAZ



Population intervention effect among studies with -1 to -0.5 mean LAZ



Population intervention effect among studies with -0.5+ mean LAZ



References

1. Young JG, Herńan MA, Robins JM. Identification, estimation and approximation of risk under interventions that depend on the natural value of treatment using observational data. *Epidemiol Methods*. 2014 Dec;**3**(1):1–19.

2. Westreich D. From Patients to Policy: Population Intervention Effects in Epidemiology. *Epidemiology*. 2017 Jul;**28**(4):525–528.